

文章编号: 100120920(2008)03-02632-04

# 一类分层非结构化 P2P 系统的随机切换模型

徐陈锋, 奚宏生, 江琦, 殷保群

(中国科学技术大学 自动化系, 合肥 230027)

**摘要:** 对于一类利用集中式构架和分布式构架各自优点的分层非结构化 P2P 系统, 通过定义一种 Markov 切换空间模型来描述其动态分组切换行为. 在 Markov 决策过程理论的基础上, 给出了关于性能指标的策略迭代和在线策略迭代算法, 并通过实例仿真说明该方法的优越性.

**关键词:** Markov 决策过程; 分层非结构化 P2P; 超级节点; 策略迭代

**中图分类号:** O232      **文献标识码:** A

## Stochastic switching model for a class of hierarchical unstructured P2P systems

XU Chenfeng, XI Hongsheng, JIANG Qi, YIN Baohun

(Department of Automation, University of Science and Technology of China, Hefei 230027, China. Correspondent: XU Chenfeng, E-mail: johnx@mail.ustc.edu.cn)

**Abstract:** For a class of hierarchical unstructured peer-to-peer (P2P) systems, a Markov switching space model is introduced to describe their behavior of dynamic grouping. Based on the theory of Markov decision processes, the policy iteration and online policy iteration algorithms for performance optimization are provided. And a simulation example shows the advantage of the proposed method.

**Key words:** Markov decision process; Hierarchical unstructured P2P; Super peer; Policy iteration

### 1 引言

P2P 系统的基本拓扑结构主要有集中式和分布式两种. 在集中式拓扑构架中, 节点通过中央主机进行查询, 这种方式受到中央主机性能和可靠性的制约. 非结构化 P2P 系统的查询采用泛洪的方式, 易造成网络带宽的浪费. 将上述两种机制有效地结合起来, 可以构建更为优良的系统. 具体说, 是将网络划分成多个物理拓扑相关的组, 每一组选择一个超级节点, 该超级节点储存了本组内资源的索引信息. 各节点通过本组的超级节点进行查询, 若超级节点上有所需的信息, 则返回给查询节点; 否则, 由超级节点以泛洪的方式向其他超级节点查询, 最终将所需信息返回给查询节点.

对于这种类型的 P2P 网络, 组的数量和大小 (即组中所含节点数) 等因素, 对网络的查询效率和维持网络的各种开销都有一定的影响. 从直观上看,

如果组的数量很多而每组的节点数很少, 则组间查询会很多, 传输代价就高; 反之, 过大的组使得超级节点维护本组的开销很大, 会影响其性能和可靠性. P2P 网络的节点分布和节点多少是随时间动态变化的. 为了平衡不同的指标, 要求系统能对组进行动态调整. 对于一个确定的 P2P 网络, 节点到达 (上线)、逗留和离开 (离线), 可用一种动态过程来描述. 系统对组的动态调整构成了不同动态过程之间的切换. 对这种切换行为进行描述, 结合动态过程构成系统模型, 便可分析和优化系统.

在相关的文献中, P2P 系统的性能一般由仿真结果给出<sup>[1,2]</sup>. 文献[3]采用 Markov 链描述结构化 P2P 系统的查询, 以此来分析系统的弹性; [4]提供了一套基于 Markov 链的 P2P 性能模型构建方法, 用来分析查询的效率; [5]给出了一个分析模型, 用来比较非结构化 P2P 系统的泛洪搜索策略. 以上模

收稿日期: 2007-12-24; 修回日期: 2008-02-27.

基金项目: 国家自然科学基金项目(60574065, 60774038); 国家 863 计划项目(2006AA01Z114); 中国下一代互联网示范工程项目(CNGI04152A); 安徽省自然科学基金项目(050420301, 070412063).

作者简介: 徐陈锋(1980), 男, 浙江海宁人, 博士生, 从事离散事件动态系统等研究; 奚宏生(1950), 男, 上海人, 教授, 博士生导师, 从事离散事件动态系统、通信网络等研究.

型化方法中,并未涉及或直接涉及具有超级节点的非结构化 P2P 分层网络,即未考虑分层结构因素对性能的影响.文献[1]比较了具有超级节点的非结构化 P2P 分层网络与普通的非结构化 P2P 网络,前者具有巨大的优势,但未讨论分层结构的差异对各种性能的影响,也没有对问题模型化.

本文利用 Markov 切换空间模型来描述这类 P2P 网络的动态分组切换行为.在该类 P2P 网络的每个分组模式下,状态过程对应于一个状态空间,其动态特性可用 Markov 过程来描述;在给定的多个分组模式下,系统的动态特性可通过多个状态空间之间的切换所组成的 Markov 切换空间过程来描述.在此基础上,将策略迭代和在线的策略迭代算法引入模型,优化这类分层非结构化 P2P 网络的性能.

## 2 系统模型

### 2.1 一类分层非结构化 P2P 系统的动态分组问题

相对于基于分布式散列表的结构化 P2P 系统,基于泛洪式搜索的非结构化 P2P 网络的扩展性和容错性较好,但控制信息的泛滥消耗了大量带宽,并可能造成网络拥塞,具有很高的网络传输代价.将整个非结构化 P2P 网络分成多个由若干节点组成的组,每个组选出一个超级节点,用于管理所在的组,缓存本组内资源的索引信息以及部分近期的搜索结果.各节点通过本组的超级节点进行查询,若超级节点上有所需的信息,则返回给查询节点;否则,由超级节点以泛洪的方式向其他超级节点查询,最终将所需信息返回给查询节点.这种作法可在保证搜索效率和成功率的前提下,大大减少网络的传输代价.实际上,这种作法将网络分成了两层:第一层是组内系统,其管理和搜索都以超级节点为中心,是偏集中式的结构;第二层是由以超级节点为代表的组间系统,采用纯分布式 P2P 方式和基于泛洪的搜索.

对于这类网络而言,在某个确定的时间段内,如果一个组的规模很小而组很多,则在组间进行泛洪式搜索的几率很大,不能有效地减少搜索的网络传输代价;而一个规模过大的组,将影响超级节点的性能,甚至超过其负载能力,且这一层的管理和搜索都以超级节点为中心,易造成系统的不稳定.另外,系统是随时间变化的,系统的规模和节点分布会发生变化,使得几个相邻的较小的组需要合并,而较大的组需要分裂.即系统需要由一个分组结构切换到另一个分组结构,以保证系统的总体性能.

考虑以路由的跳跃点数和网段划分等物理网络拓扑相关的因素为基本分组原则.为了降低复杂性,将整个 P2P 系统的覆盖区域分成一个个不可再分

的最小区域.最小区域中节点之间的通信不需经过路由器,它们处于相同且规模较小的网段.组就是由一个或多个物理上相邻的最小区域构成的.

**定义 1 最小区域**是处于网络边缘且物理上相邻,相互通信不需经过路由器,处于相同且规模较小的网段的节点构成的逻辑单位,是人为规定的构成一个组的最小单位.

**定义 2 分组结构**是对最小区域的一种划分,划分在一起的最小区域构成一个组,并且这些划分在一起的最小区域总是物理上相邻的.这里所说的物理上相邻,是指这些构成一个组的最小区域之间的通信经过较少的路由器.

**定义 3 分组切换**是在系统运行过程中对最小区域的划分发生的变化,即系统分组结构的改变.从现象上看,就是由一个组分裂为两个组,或由两个组合并为一个组.

假设这些最小区域不能同时属于两个或两个以上的组.对于一个确定的物理网络,分组结构是与物理网络相关的,因此合理的分组结构数是有限的.

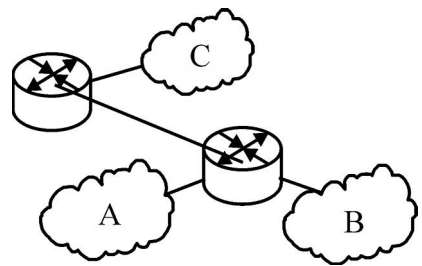


图 1 示例网络

考虑具有 3 个不可再分的最小区域 A, B 和 C 的 P2P 网络,如图 1 所示.在路由上, A 与 B 中间隔了一个路由器, A, B 与 C 隔了两个路由器.据此可给出几种较为合理的划分,例如:

- 1) {A, B, C};
- 2) {A, B}, {C};
- 3) {A}, {B}, {C}.

假设在某个时刻,分组切换只能由两个组变为一个组,或由一个组变为两个组.则示例网络中的分组结构 1) 与分组结构 3) 之间不能直接切换,而 1) 与 2), 2) 与 3) 能够互相直接切换.

### 2.2 随机切换模型

考虑一个分层非结构化 P2P 网络.首先规定它的最小区域,并且考虑这些最小区域可能的合理划分,每种划分就是一个分组结构.设它有  $K$  个合理的分组结构,其集合为  $M = \{1, 2, \dots, K\}$ .在当前分组结构  $k \in M$  下,网络被分成  $m_k$  组,其中第  $i$  组的节点个数为  $n_i^k, n_i^k \in [N_i^k, N_i^k]$  为组  $i$  中最大可能的 P2P 节点数.

假设整个网络中只存在单一种类的节点,组  $i$

中节点的到达率为  $K_i^k$ , 每个节点的在线时间服从均值为  $1/L$  的指数分布. 则在分组结构  $k \in M$  下, 系统运行的状态子过程  $X_t^k$  是定义在有限状态空间  $S_k = \{s_k = (n_1^k, n_2^k, \dots, n_{m_k}^k)\}$  上的 Markov 过程<sup>[6]</sup>, 它描述了各组的节点数变化. 其转移速率矩阵  $A_k = [a_{ij}^k]$  为  $L_k \times L_k$  阶矩阵, 其中  $L_k = \sum_{i=1}^{m_k} (N_i^k + 1)$ .  $A_k$  可由所有组的转移速率矩阵获得, 设  $A_i^k = [a_{pq}^{ki}] (i = 1, 2, \dots, m_k)$  为各组的转移速率矩阵. 根据相关排队理论<sup>[7]</sup>, 有

$$a_{pq}^{ki} = \begin{cases} K_i, & q = p + 1; \\ L_i, & q = p - 1; \\ -(K_i + L_i), & p = q; \\ 0, & \text{其他.} \end{cases} \quad (1)$$

其中  $K_i$  和  $L_i$  分别为组  $i$  中节点数为  $p$  时的节点到达速率和节点离开速率.

定义切换行动为  $u^k$ , 表示从分组结构为  $k \in M$  切换到分组结构  $kc \in M$  的行动. 表现在物理上, 是由两个组合并为一个组, 或由一个组分裂为两个组, 从而构成了分组切换. 当  $k = kc$  时, 系统不进行分组切换.

定义一个 Markov 切换空间过程  $X_t$ , 其状态空间为  $S = \{(k, s_k), s_k \in S_k, k \in M\}$ . 考虑确定平稳策略  $S = (d, d, \dots) = d^1$ , 其中  $d$  为决策规则. 即  $d: (k, s_k) | y u^k, k, kc \in M, u^k$  为切换行动. 则过程  $X_t$  仍是一个 Markov 过程. 设  $O$  为确定平稳策略构成的集合, 在平稳策略  $S \in O$  下, 从状态  $s_k$  转移到  $s_{kc}$  的概率为

$$p_{s_k s_{kc}}^S = \begin{cases} \sum_{\substack{k' \in M \\ s_{k'} \in S_{k'}}} p_{s_k s_{k'}}^k p_{s_{k'} s_{kc}}^{k'}, & d(k, s_k) = u^k, k \neq kc; \\ p_{s_k s_{kc}}^k, & d(k, s_k) = u^k, k = kc; \\ 0, & \text{其他.} \end{cases} \quad (2)$$

根据状态转移速率矩阵的定义, 可得  $A^S = [a_{s_k s_{kc}}^S]$ , 其中

$$a_{s_k s_{kc}}^S = \begin{cases} \sum_{\substack{k' \in M \\ s_{k'} \in S_{k'}}} a_{s_k s_{k'}}^k p_{s_{k'} s_{kc}}^{k'}, & d(k, s_k) = u^k, k \neq kc; \\ a_{s_k s_{kc}}^k, & d(k, s_k) = u^k, k = kc; \\ 0, & \text{其他.} \end{cases} \quad (3)$$

式中  $p_{s_k s_{kc}}^k = P(s_{kc} | s_k)$ , 为当状态进入  $s_k$  并切换到  $kc (kc \in M)$  时状态变为  $s_{kc}$  的概率.

定义性能代价函数  $f \in B(S) \times R$ , 在策略  $S \in O$  下, 系统的平均性能测度为

$$G^S = \lim_{T \rightarrow \infty} \frac{1}{T} E \int_0^T f(X_t) dt. \quad (4)$$

综合上述, 该切换模型可表示为  $\{S, A^S, f, G^S\}$ , 而动态分组切换问题就是寻找一个策略  $S^*$ , 使得平均测度最小, 即  $S^* \in \arg\{\min_{S \in O} G^S\}$ .

### 2.3 策略迭代法寻优

本节在文献[8, 9] 的基础上, 给出 Markov 切换空间过程关于性能指标的优化算法.

根据相关性势理论<sup>[10]</sup>, 如果在策略  $S$  下  $X_t$  是遍历的, 则过程有唯一的稳态分布  $p^S$ , 满足  $p^S e = 1, p^S A^S = 0$ . 则性能势

$$g^S = -(A^S)^{\#} f + B, \quad B \in R \quad (5)$$

为 Poisson 方程  $A^S g^S + f = G^S e$  的解, 其中

$$(A^S)^{\#} = (A^S + \epsilon p^S)^{-1} - p^S \quad (6)$$

为  $A^S$  的群逆. 对于每个  $B \in R$ , 式(5) 都是 Poisson 方程的解, 可令  $B = 0$ . 于是, 系统的最优性方程为

$$A^S g^{S^*} + [A^S g^{S^*} + P] S \in O. \quad (7)$$

基于此最优性方程, 根据策略优化算法<sup>[8]</sup> 可得理论的最优策略. 但该算法的性能势由式(5) 直接计算, 这对于大状态空间问题计算复杂度很高. 可考虑基于样本轨道估计性能势<sup>[11, 12]</sup>, 以减少相关矩阵的运算. 构建在线的策略迭代优化算法<sup>[9]</sup> 如下:

1) 任选  $S^0 \in O$ , 令  $k = 0, g^{S^0} = 0$ , 选定  $N \in (0, 1/2), C > 0$ .

2) 观测  $S^k \in O$  下的系统, 计算  $V_k(s^*, \mathbf{x})$ . 如果  $\mathbf{x} = s^*$ , 则  $V_k(s^*, \mathbf{x}) = 0$ ; 否则

$$V_k(s^*, \mathbf{x}) = \begin{cases} \int_0^{l_{k-1}(s^*)} f(X_t) dt, & V(\mathbf{x}) = 1; \\ g^{S^{k-1}}, & V(\mathbf{x}) = 0. \end{cases} \quad (8)$$

其中:  $s^*$  为选定的更新状态;  $l_k(s^*)$  为  $X_t$  的更新时间, 满足

$$\begin{cases} l_0(s^*) = 0, k \geq 1, \\ l_k(s^*) = \inf\{t \in T : X_t = s^*\}, \\ l_k(\mathbf{x}) = \inf\{t \in T : X_t = \mathbf{x}\}. \end{cases} \quad (9)$$

当且仅当  $l_k(\mathbf{x}) < l_{k-1}(s^*)$  时, 有  $V(\mathbf{x}) = 1$ . 另外有

$$G^S = \frac{\int_0^{l_{k+1}(s^*)} f(X_t) dt}{l_{k+1}(s^*) - l_k(s^*)}. \quad (10)$$

更新性能势

$$g^{S^k}(\mathbf{x}) = g^{S^{k-1}}(\mathbf{x}) + \frac{1}{k+1} [V_k(s^*, \mathbf{x}) - g^{S^{k-1}}(\mathbf{x})]. \quad (11)$$

3) 求解  $A \in O$ , 使得对于每个状态  $s \in S, \sum_{s' \in S} a_{ss'}^A g^{S^k}(s')$  最小. 如果存在状态  $s \in S$ , 使得

$$\sum_{s' \in S} a_{ss'}^A g^{S^k}(s') \leq \frac{C}{(k+1)^{1/2-N}}, \quad (12)$$

则令  $S^{k+1} = A$  否则,  $S^{k+1} = S^k$ .

4) 令  $k = k + 1$ , 返回第 2) 步.

### 3 应用仿真

考虑图 1 有 3 个不可再分的最小区域 A, B 和 C 的 P2P 系统, 设其合理的分层结构有 3 个, 分别为: 1) {A, B, C}; 2) {A, B}, {C}; 3) {A}, {B}, {C}.

设在某个时刻发生分组切换的情况只能同时由两组变为一组, 或由一组变为两组, 则分组结构 1) 和分组结构 3) 之间不能直接切换. 设区域 A, B 和 C 的最大 P2P 节点数分别为  $N_a, N_b$  和  $N_c$ , 状态空间 S 中的状态根据字典序排列, 并且假定这些节点只处于两种状况: 在 P2P 系统中和在不在 P2P 系统中. 设节点不在 P2P 系统中的时间服从均值为  $1/K$  的指数分布, 则在分组结构  $k \in M$  下, 第  $i$  组的节点到达率  $K_i^k$  满足  $K_i^k = (N_i^k - n_i^k)K$ ; 设节点在 P2P 系统中的时间服从均值为  $1/L$  的指数分布, 则 P2P 系统在分组结构  $k \in M$  下, 第  $i$  组中节点离开系统的速率为  $L_i^k = n_i^k L$ .

定义性能函数  $f(s) = A f_1(s) + B f_2(s)$  为代价  $f_1(s)$  和  $f_2(s)$  的加权, 其中 A 和 B 为加权系数. 设每个超级节点维护本组信息和邻居信息的代价

$$f_1(s) = \begin{cases} a n_1, & k = 1; \\ \frac{1}{2} a (n_1 + n_2) + b, & k = 2; \\ \frac{1}{3} a (n_1 + n_2 + n_3) + 2b, & k = 3. \end{cases} \quad (13)$$

其中:  $s = (k, \mathbf{s})$ ,  $a$  为维护本组内节点信息的平均代价,  $b$  为维护邻居信息的平均代价. 组的规模越大, 超级节点维护的信息越多, 代价越大; 组的数目越多, 超级节点之间保持相互联络的代价越大.

设平均搜索步长

$$f_2(s) = \begin{cases} 1, & k = 1; \\ 1 + \frac{1}{2} \sum_{i=1}^2 (1-p)^{n_i}, & k = 2; \\ 1 + \frac{1}{3} \sum_{i=1}^3 (1-p)^{n_i}, & k = 3. \end{cases} \quad (14)$$

其中:  $s = (k, \mathbf{s})$ ,  $p$  为在各个节点上所需资源(如共享文件)存在的平均概率. 搜索时, 首先通过本地的超级节点查询本组, 如果查不到, 则通过泛洪搜索在组间搜索. 平均搜索步长代表搜索的效率, 步长越短, 搜索越快. 如果组间的泛洪搜索使用随机漫步, 则可代表搜索的代价, 步长越短, 代价越小.

仿真参数选取  $K = L = 1, A = B = 1, a = b = 1, N_a = N_b = N_c = 3, p = 0.01$ ; 在线策略迭代算法选取误差控制参数  $C = 0.01, N = 0.002$ .

不进行动态分组切换, 系统始终使用某一固定分层结构的平均代价如表 1 所示; 分别利用策略迭代算法<sup>[8]</sup> 和在线策略迭代算法<sup>[9]</sup> 的优化策略, 进行动态分组切换时系统的平均代价和相应的算法计算效率如表 2 所示. 采用优化动态分组切换的方式, 比采用固定分层结构方式可取得更小的性能代价; 使用在线策略迭代算法, 性能势是在一条样本轨道上估计得到的, 无需进行式(5) 的矩阵运算, 因此计算效率得到显著的提高. 由于存在估计误差, 所得结果无法达到理论最优值, 如表 2 所示.

表 1 固定分层结构下的性能

固定分层结构	平均代价
$k = 1$	5.500
$k = 2$	5.228
$k = 3$	5.485

表 2 动态分组切换下的性能

优化算法	平均代价	计算耗时 / s
策略迭代	5.034	173.4
在线策略迭代	5.067	86.6

### 4 结 语

本文分析了分层结构差异对系统性能的影响, 对一类分层非结构化 P2P 系统动态分组切换的行为进行模型化, 给出了相应的优化问题, 并通过一个例子说明优化策略的效果. 优化的动态分组切换可根据系统的变化调整系统, 具有更好的性能; 与直接计算性能势相比, 基于样本轨道估计性能势可提高算法的效率. 按照给出的模型, 实际系统的状态空间很大, 不可避免地会存在维数灾问题. 下一步工作将考虑使用状态聚集、参数化梯度算法等方法, 进一步降低单步迭代的计算复杂度, 提高算法的实时性.

### 参考文献(References)

[1] Benevenuto F, Ismael J J, Almeida J. Quantitative evaluation of unstructured peer2topeer architectures [C]. Proc of the 2004 Int Workshop on Hot Topics in Peer2toPeer Systems. Washington: IEEE Computer Society, 2004: 50-65.

[2] Lin T N, Wang H P, Wang J M. Search performance analysis and robust search algorithm in unstructured peer2topeer networks[C]. Proc of 2004 IEEE Int Symp on Cluster Computing and the Grid. Piscataway: IEEE Press, 2004: 346-354.

[3] Wang S Q, Xuan D, Zhao W. Analyzing and enhancing the resilience of structured peer2topeer systems[J]. J of Parallel and Distributed Computing, 2005, 65 (2): 202-219.

- [8] Kwak K C, Pedrycz W. Face recognition: A study in information fusion using fuzzy integral [J]. *Pattern Recognition Letters*, 2005, 26(6): 7192733.
- [9] 李德清, 崔红梅, 李洪兴. 基于层次变权的多因素决策 [J]. *系统工程学报*, 2004, 19(3): 258263.  
(Li Deqing, Cui Hongmei, Li Hongxing. Multifactor decision making based on hierarchical variable weight [J]. *J of Systems Engineering*, 2004, 19(3): 258263.)
- [10] 李德清, 李洪兴. 变权决策中变权效果分析与状态变权向量的确定 [J]. *控制与决策*, 2004, 19(11): 1241245.  
(Li Deqing, Li Hongxing. Analysis of variable weights effect and selection of appropriate state variable weights vector in decision making [J]. *Control and Decision*, 2004, 19(11): 1241245.)
- [11] Ramanathan R. Data envelopment analysis for weight derivation and aggregation in the analytic hierarchy process [J]. *Computers and Operations Research*, 2006, 33(5): 128921307.
- [12] 王建军, 杨德礼. 供应商选择的 AHP/Promethee 决策方法 [J]. *管理评论*, 2006, 18(7): 5261.  
(Wang Jianjun, Yang Deli. An AHP/Promethee based method of selecting supplier [J]. *Management Review*, 2006, 18(7): 5261.)
- [13] 骆正清. AHP 中不一致性判断矩阵调整的新方法 [J]. *系统工程理论与实践*, 2004, 24(6): 8492.  
(Luo Zhengqing. A new method for adjusting inconsistency judgment matrix in AHP systems [J]. *Systems Engineering*) Theory and Practice, 2004, 24(6): 8492.)
- [14] 朱建军, 刘士新, 王梦光. 一种新的改进不一致判断矩阵的方法 [J]. *系统工程理论与实践*, 2003, 23(11): 9998, 110.  
(Zhu Jianjun, Liu Shixin, Wang Mengguang. A new algorithm to improve the inconsistent comparison matrix systems [J]. *System Engineering*) Theory and Practice, 2003, 23(11): 9998, 110.)
- [15] Kojadinovic I, Marichal J L, Roubens M. An axiomatic approach to the definition of the entropy of a discrete Choquet capacity [J]. *Information Sciences*, 2005, 172(1/2): 132153.
- [16] Marichal J L. Entropy of discrete Choquet capacities [J]. *European J of Operational Research*, 2002, 137(3): 612624.
- [17] 朱勇珍, 李洪兴. 状态变权的公理化体系和均衡函数的构造 [J]. *系统工程理论与实践*, 1999, 19(7): 112118, 131.  
(Zhu Yongzhen, Li Hongxing. Axiomatic system of state variable weights and construction of balance functions [J]. *Systems Engineering*) Theory and Practice, 1999, 19(7): 112118, 131.)
- [18] Chiang J H. Choquet fuzzy integral based hierarchical networks for decision analysis [J]. *IEEE Trans on Fuzzy Systems*, 1999, 7(1): 6371.

(上接第 266 页)

- [4] Spognardi A, Di Pietro R. A formal framework for the performance analysis of P2P networks protocols. <http://www.cecs.uci.edu/papers/ipdps06/pdfs/1568976643HOTP22paper21.pdf>, 2006.
- [5] Gaeta R, Sereno M. Model based evaluation of search strategies in peer-to-peer networks. <http://www.cecs.uci.edu/papers/ipdps06/pdfs/1568976692HOTP22pape21.pdf>, 2006.
- [6] Puterman M L. Markov decision processes: Discrete stochastic dynamic programming [M]. New York: Wiley, 1994.
- [7] Cooper P R. Introduction to querying theory [M]. New York: Elsevier North Holland, 1981.
- [8] Tang H, Xi H S, Yin B Q. Performance optimization of continuous-time Markov control processes based on performance potentials [J]. *Int J of Systems Science*, 2003, 34(1): 63271.
- [9] Fang H T, Cao X R. Potential based online policy iteration algorithms for Markov decision processes [J]. *IEEE Trans on Automatic Control*, 2004, 49(4): 492505.
- [10] Cao X R, Chen H F. Perturbation realization, potentials and sensitivity analysis of Markov processes [J]. *IEEE Trans on Automatic Control*, 1997, 42(10): 138221393.
- [11] Cao X R. The potential structure of sample paths and performance sensitivities of Markov systems [J]. *IEEE Trans on Automatic Control*, 2004, 49(12): 21222142.
- [12] Chong E K P, Ramadge P J. Stochastic optimization of regenerative systems using infinitesimal perturbation analysis [J]. *IEEE Trans on Automatic Control*, 1994, 39(7): 14021410.